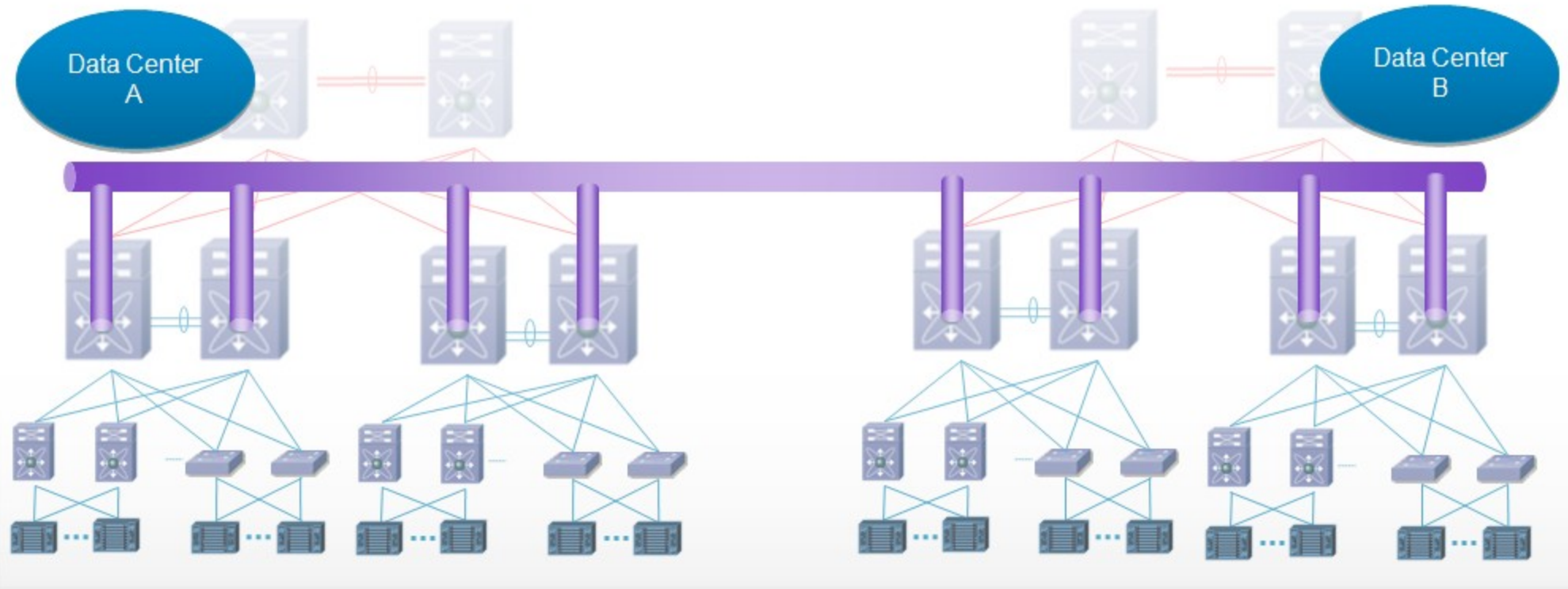
OTV
Overlay Transport Virtualization

Lauri Toropainen

ltoropai@cisco.com

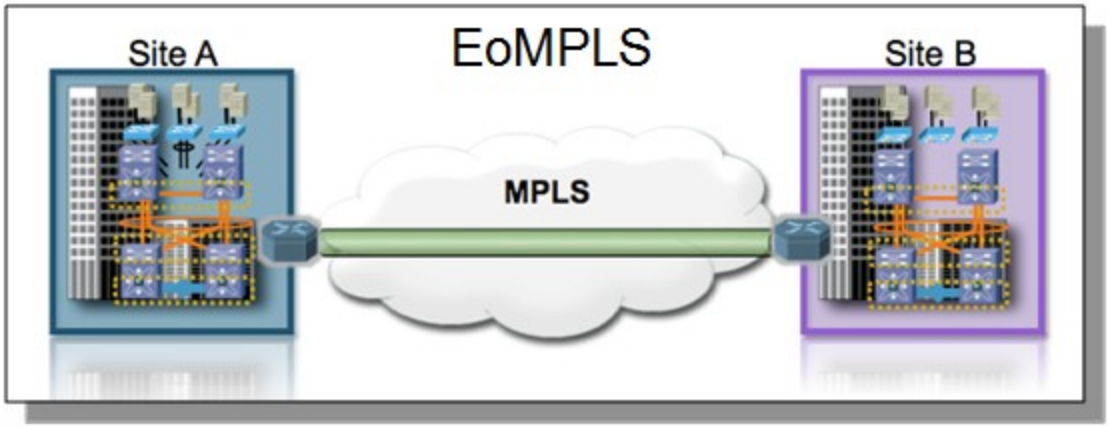
Problem Statement
LAN extensions at Layer 2: Inter/intra Data Center

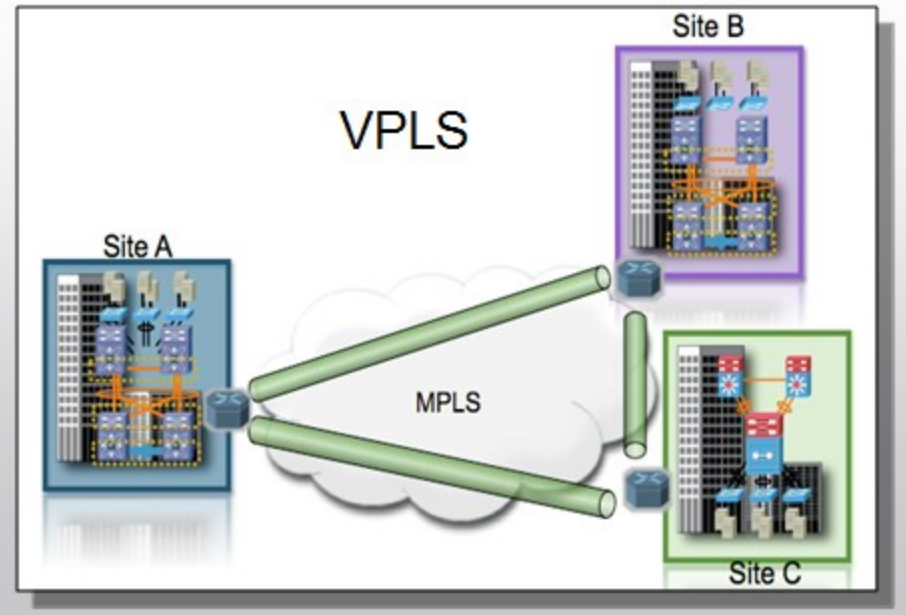


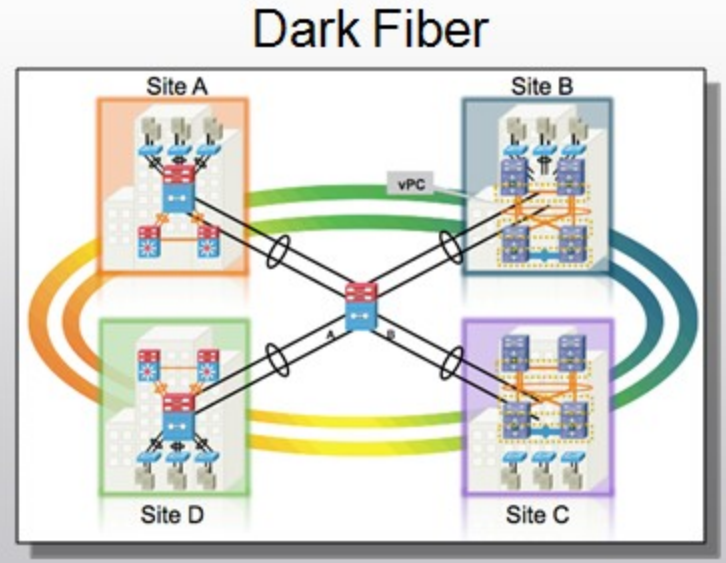
- *Business Needs*
  - Disaster Avoidance
  - Business Continuance
  - Workload mobility (server virtualization)

- *IT Solutions*
  - Active/Standby Migration
  - Server HA clusters, "Geo-clustering"
  - Move, consolidate servers, "Vmotion"

Traditional Layer 2 VPNs

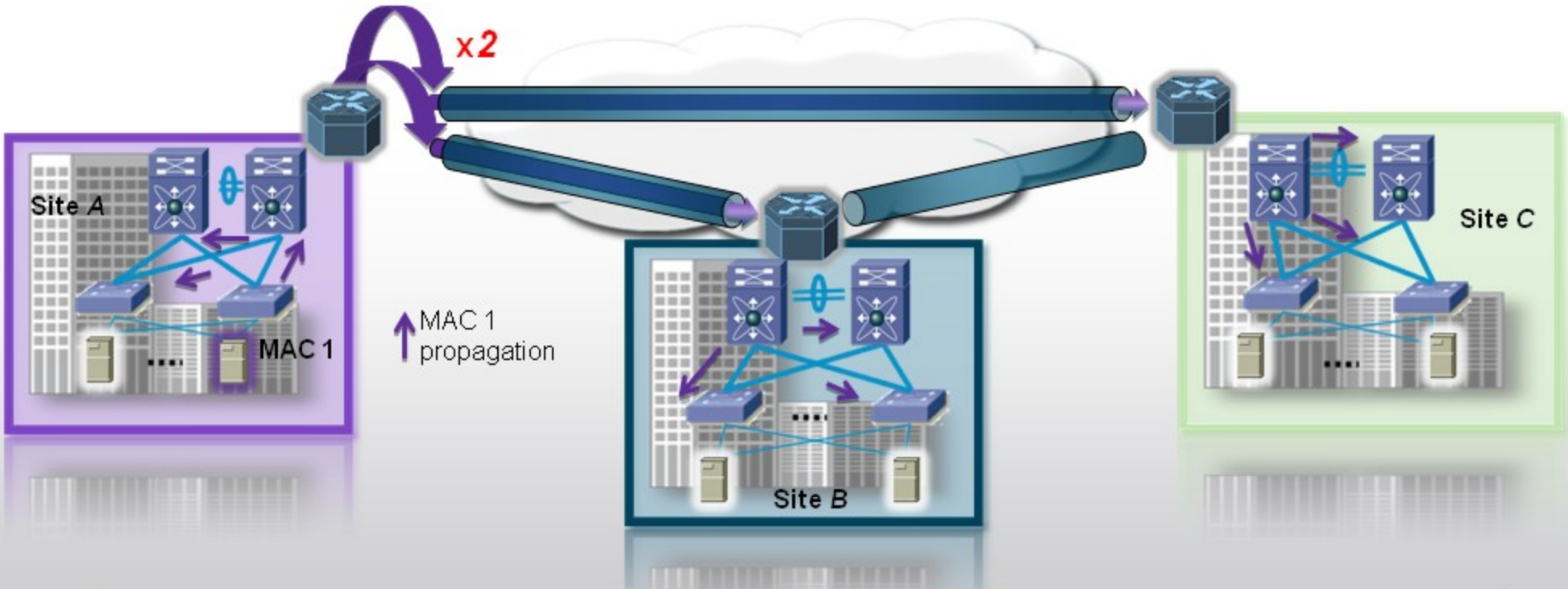






Flooding Behavior

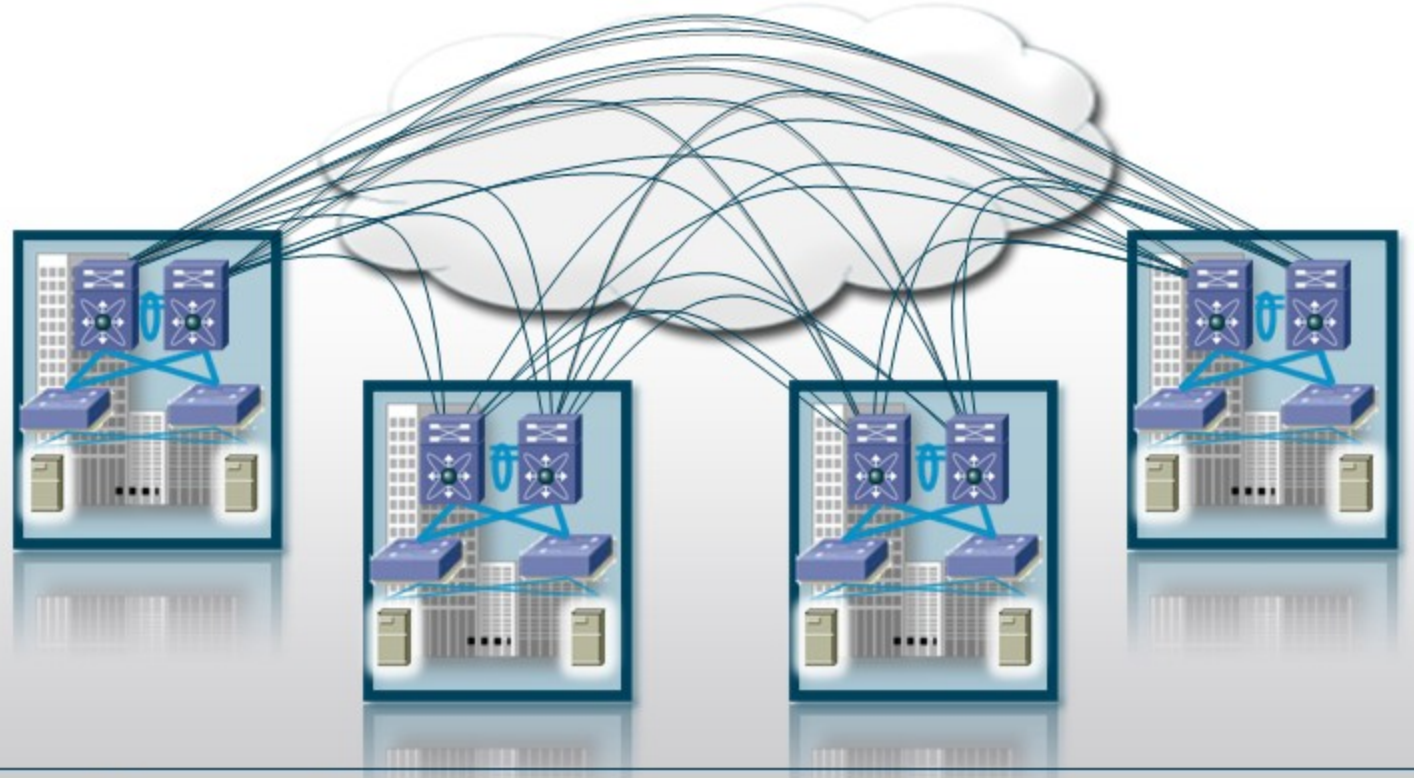
- Traditional Layer 2 VPN technologies rely on flooding to propagate MAC reachability.
- The flooding behavior causes failures to propagate to every site in the L2-VPN.



- A solution that provides layer 2 connectivity, yet restricts the reach of the flood domain is necessary in order to contain failures and preserve the resiliency.

Pseudo-wires Maintenance

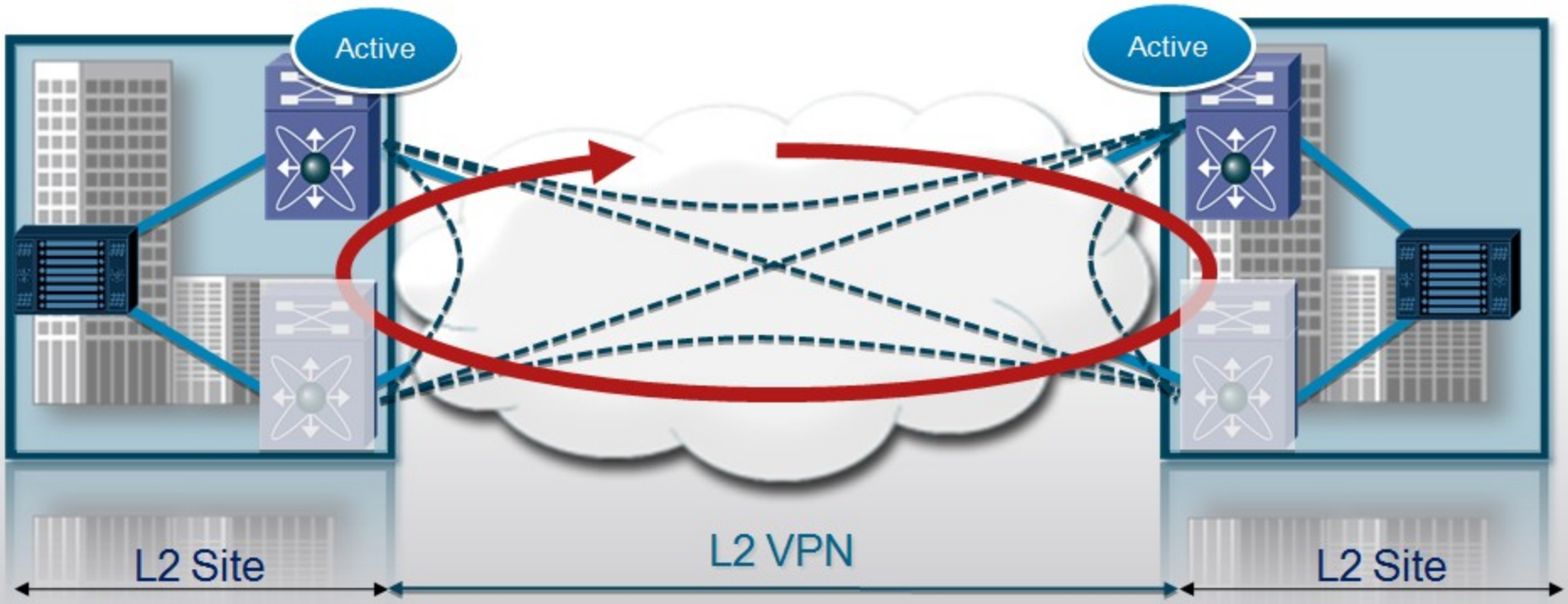
- Before any learning can happen a full mesh of pseudo-wires/tunnels must be in place.
- For N sites, there will be N*(N-1)/2 pseudo-wires. Complex to add/remove sites.
- Head-end replication for multicast and broadcast. Sub-optimal BW utilization.



- A simple overlay protocol with built-in functionality and point-to-cloud provisioning is key to reducing the cost of providing this connectivity

Multi-Homing

- Require additional protocols to support Multi-homing.
- STP is often extended across the sites of the Layer 2 VPN. Very difficult to manage as the number of sites grows.
- Malfunctions on one site will likely impact all sites on the VPN.



- A solution that natively provides automatic detection of multi-homing without the need of extending the STP domains is key.

What can be improved

- Data Plane Learning → Control Plane Learning

  Moving to a Control Plane protocol that proactively advertises MAC addresses and their reachability instead of the current flooding mechanism.

- Pseudo-wires and Tunnels → Dynamic Encapsulation

  No static tunnel or pseudo-wire configuration required.

  Optimal replication of traffic done closer to the destination, which translates into much more efficient bandwidth utilization in the core.

- Multi-Homing → Native Built-in Multi-homing

  Ideally a multi-homed solution should allow load balancing of flows within a single VLAN across the active devices in the same site, while preserving the independence of the sites.

  STP confined within the site (each site with its own STP Root bridge)

Overlay Transport Virtualization

Technology Pillars

OTV is a "MAC in IP" technique for supporting Layer 2 VPNs **OVER ANY TRANSPORT**.

Dynamic Encapsulation

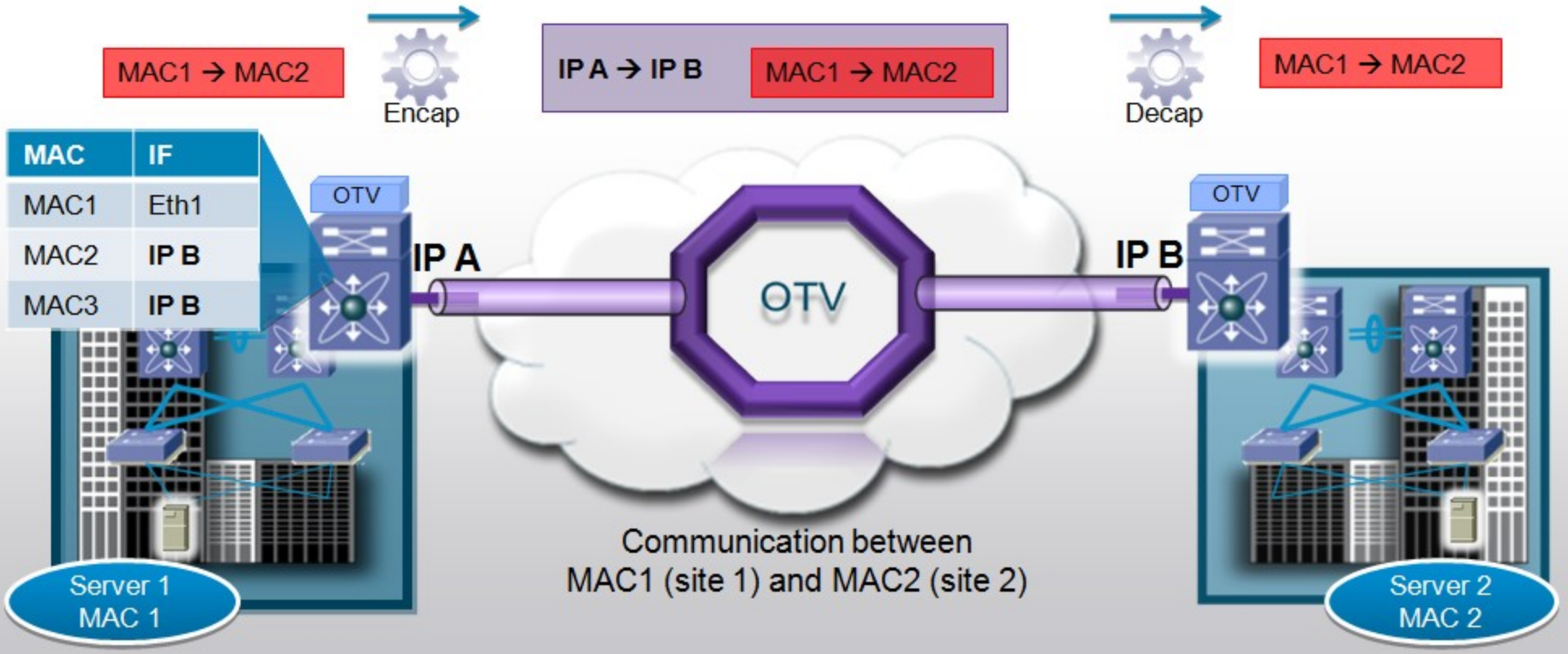
- No Pseudo-Wire State Maintenance
- Optimal Multicast Replication
- Multi-point Connectivity
- Point-to-Cloud Model

Protocol Learning

- Built-in Loop Prevention
- Preserve Failure Boundary
- Seamless Site Addition/Removal
- Automated Multi-homing

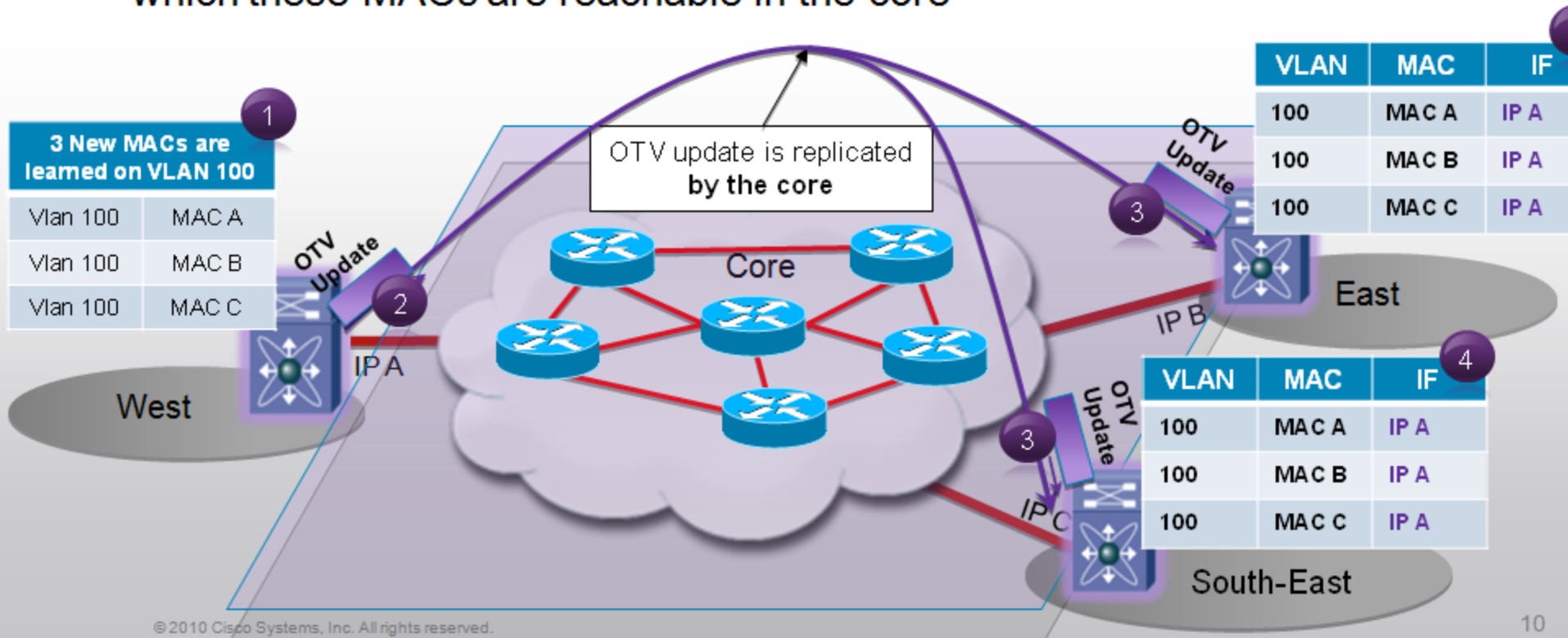
OTV at a Glance

- Ethernet traffic between sites is encapsulated in IP: “MAC in IP”
- Dynamic encapsulation based on MAC routing table
- No Pseudo-Wire or Tunnel state maintained



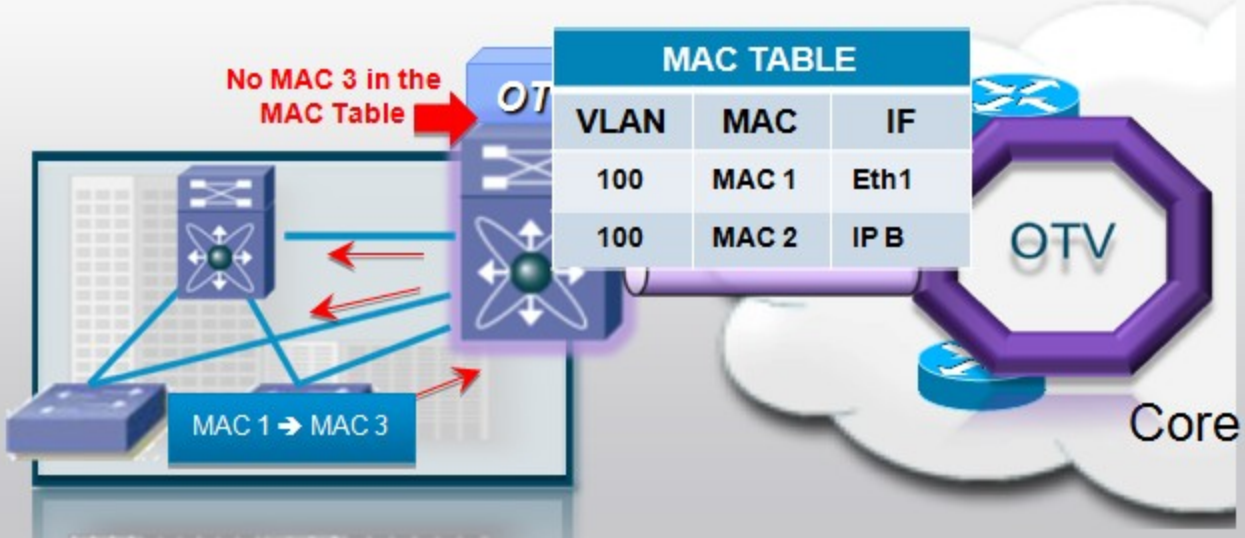
Building the MAC tables – Multicast

- The OTV control plane **proactively advertises** MAC reachability to other sites by using **multicast**
- Every time an Edge Device learns a new MAC address, the OTV control plane will advertise it together with its associated **VLAN IDs and IP next hop**
- The IP next hops are the addresses of the Edge Devices through which these MACs are reachable in the core



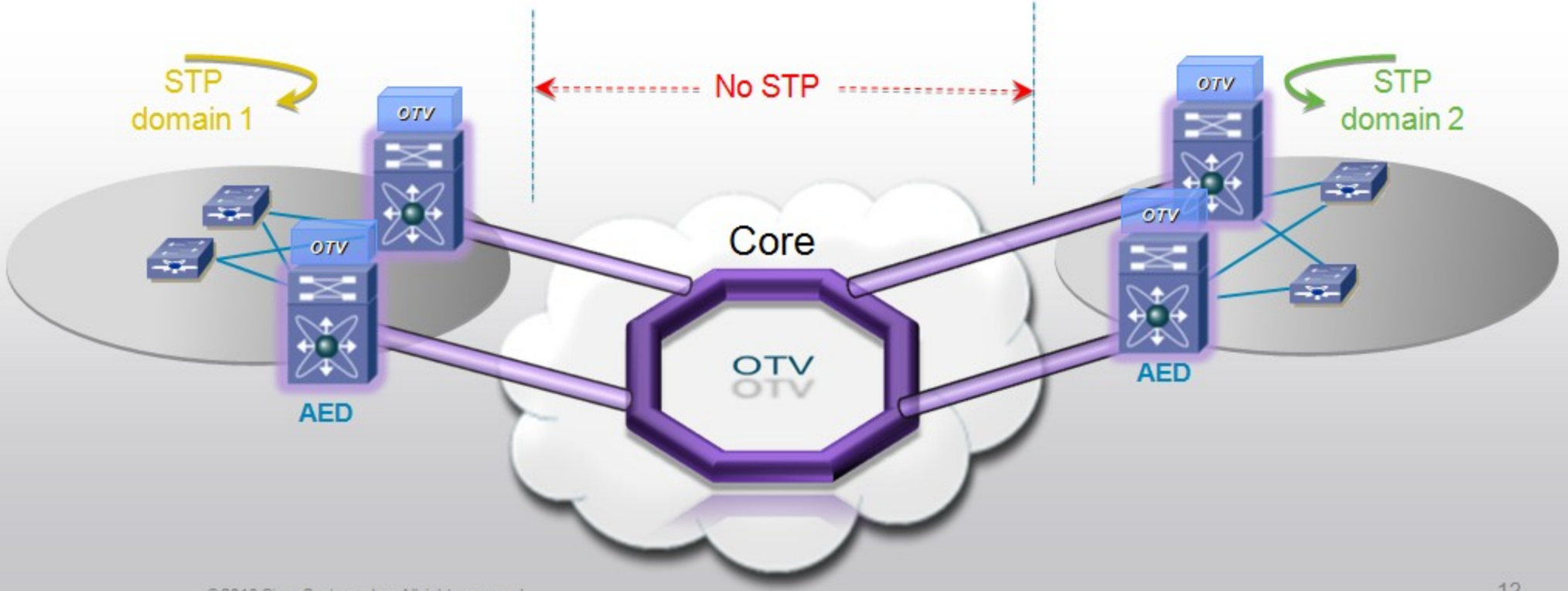
Unknown Unicast Packet Handling

- Flooding of unknown unicast over the overlay **is not required** and is therefore suppressed.
- Any unknown unicasts that reach the OTV edge device will not be forwarded onto the overlay.
- The assumption here is that the end-points connected to the network are not silent or uni-directional.
- MAC addresses for uni-directional host are learnt and advertised by snooping the host's ARP reply



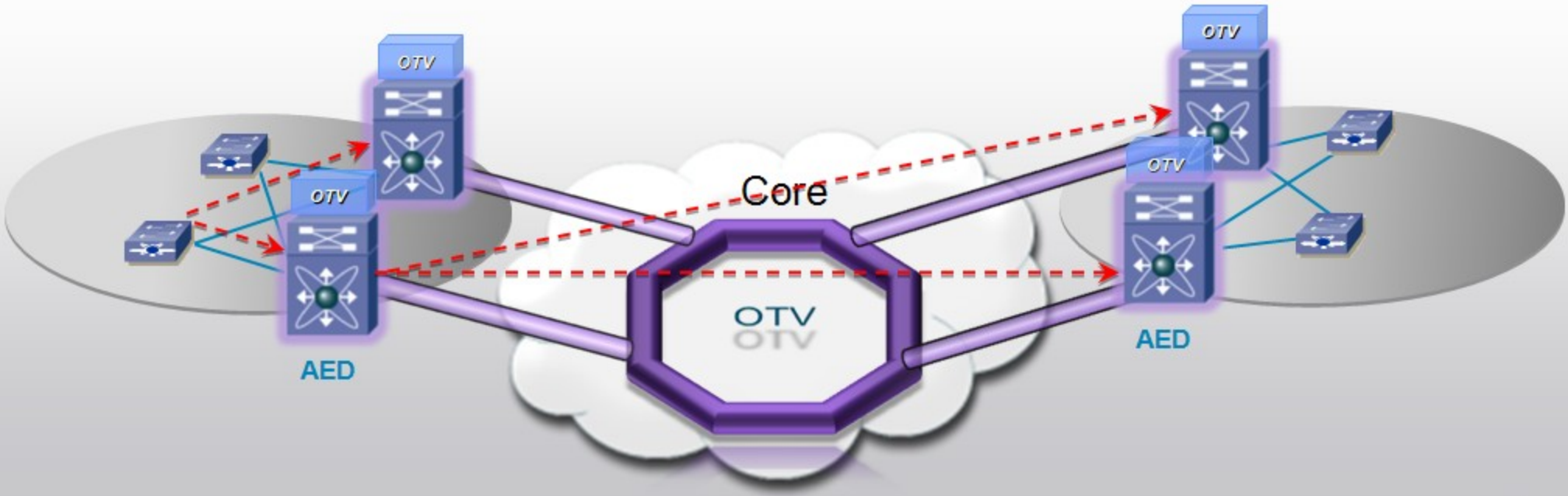
Multi-Homing & Loop Condition Handling

- OTV provides loop-free multi-homing by electing a designated forwarding device per site for each VLAN – **Authoritative Edge Device (AED)**
- OTV includes the logic necessary to avoid the creation of loops in multi-homed site scenarios – **only AEDs will forward broadcast & multicast traffic to the overlay**
- Each site will have its own STP domain, which is separate and independent from the STP domains in other sites



Active-Active ECMP & Load Balancing

- OTV allows different flows to use different edge devices when a site is multi-homed.
- The choice of the edge device is based on both the source and destination addresses of the frames to be forwarded.
- The logic affects both the election of the egress device (Load Balancing) as well as the election of the ECMP route to the remote site.



Q & A

For more information:
http://www.cisco.com/go/nexus7000